Association for Information Systems
AIS Electronic Library (AISeL)

ACIS 2005 Proceedings

Australasian (ACIS)

December 2005

Developments in e-Research: The need for a coherent IS approach

Craig McDonald
University of Canberra

Developments in e-Research: The need for a coherent IS approach

Craig McDonald

School of Information Sciences and Engineering
University of Canberra
craig.mcdonald@canberra.edu.au

Abstract

In Australia the concept of e-Research has been developing rapidly over the last three years. e-Research effort is currently focused on infrastructure building and on tools development. In the short term this will have pay-offs in large data-intensive research domains and in opportunistic niche uses of the technology. But the sustainable value of e-Research will only be realized by the broad-scale use of systems that employ that infrastructure. The Information Systems discipline has the approaches and techniques needed to make e-Research effective.

Keywords

e-Research, Human Activity System, technology-push, knowledge management, information systems

E-RESEARCH

E-Research is usually understood to be related to the use of ICT in scientific research, particularly that needing high computing power, and vast data sources in a highly distributed grid environment. Typical domains include astronomy, physics, geology and so on. The Australian Partnership for Advanced Computing (APAC) and AusGrid are typical of the kinds of organizations involved. The technologies they are developing and promoting include broadband, middleware, repositories of scientific data, sensors and instrumentation, distributed computational power, and so on. The UK e-Science Grid conceives of "an e-Scientist's workbench that supports: (I) the scientific process of experimental investigation, evidence accumulation and result assimilation; (ii) the scientist's use of the community's information; and (iii) scientific collaboration, allowing dynamic groupings to tackle emergent research problems.."

The Humanities and Social Science disciplines are increasingly engaging with e-Research. The Netherlands Institute for Scientific Information Services argues that "a number of fields have undergone radical transformations through the use of novel analytical techniques and related shifts in research paradigms. Yet, a systematic and critical interrogation of the potential of e-research paradigms and methodologies for the humanities and social sciences is hampered by disciplinary boundaries between fields, by a relative lack of resources and research infrastructures, and by the dominance of particular computational approaches in the world of e-science."

In Australia the concept of e-Research has been developing rapidly over the last three years. Both the Chair of the Australian Research Council and Australia's Chief Scientist have spoken recently about e-Science and e-Research. Cram's (2003) A Roadmap for e-Research and Batterham's (2003) E-Science: A Frontier Technology For Achieving The National Research Priorities, emphasize the use of ICT in the way research and innovation will be conducted in the future. The issues raised are well developed in the United Kingdom (eg. in the National E-Science Centre) and in the United States (eg. NSF Middleware Intitative).

The Australian Research Council launched an e-Research discussion paper (Tsoi, 2004) and, in 2005, funded a Special Research Initiative for e-Research. Later in 2005 the Australian Government created the e-Research Coordinating Committee to oversee the development of a coherent ICT infrastructure for research. The terms of reference for this Committee are summarised below:

"The virtual research environments emerging from the increasing use of distributed high-performance computing resources, data resources, grid networks and communications technologies have enabled researchers to perform their research independent of time and geographical location - interacting with colleagues, accessing remote instrumentation, sharing distributed research data and computational resources, and accessing information in digital libraries. The term "e-Research" embraces those virtual environments that facilitate real research collaborations of multidisciplinary, inter-disciplinary, or intra-disciplinary and large or small scale nature involving researchers and research organisations, nationally and internationally.

The key elements of Australia's e-Research infrastructure strategy include:

robust high-bandwidth advanced communications networks,
distributed high-performance computing and data storage capacities,
accessible data and information repositories,
accessible research instruments and facilities, and
agreed standards and coordinated middleware development." (DEST 2005)

There are, however, issues of common interests shared by all, and emerging issues which are beyond the capacity of any individual agency, institution or research organisation to address. These issues might be addressed in a more cost effective way by coordination and collaboration, nationally and internationally. For example legal issues intellectual property, cultural issues such as those related to creating trust in the 'virtual' environment; the need to provide researchers with access to the resources necessary to enable them to review the work of other scholars, access information and data in a variety of formats and to disseminate the results of their own endeavours; the need to co-ordinate activities among a number of funding initiatives in various funding agencies so that a national strategy in e-Research may be achieved; the need to engage industries and business enterprises in adopting e-Research methodology in their enterprise systems; the need to engage agencies in both federal and state governments in adopting e-Research methodology in their e-government endeavours

The question for the Information Systems discipline is what role should it be playing in e-research?

INFORMATION SYSTEMS

Information Systems (IS) is an active and interventionist discipline that mobilizes information and knowledge so that people can effectively take knowledgeable, informed actions in their organizational and social setting. It is concerned with understanding and formalizing areas of human activity and developing ICT-based systems that responsibly intervene in those areas for the benefit of all stakeholders. The benefits relate to improved metabolic processes in the system and better interaction between the system and the rest of the world with which it interacts.

Examples of some key Information Systems concepts, explored further below include Human Activity System Perspective, Task Analysis and Design, Domain Analysis and ICT-based Information and Service Architecture.

Human Activity System Perspective

Information Systems is a discipline that interests itself in the interaction of information technologies with human activity systems. *Research* can be seen as a human activity system and as such it is susceptible to investigation and intervention by Information Systems practitioners. Just as we have e-Business, e-Learning or Health Informatics, so we can have e-Research. Research as a Human Activity System interacts with the rest of the world, involves a range of people in many roles, structures of ideas and motives, and looks to the systematic use of ICT as instrumental in its activity.

Research, as a Human Activity System, can be analysed in terms of granularity from the most fine-grained, personal level of the individual researcher at work, to the broad societal level of research policy and social impact of research. Between these levels are the various organisations that mediate the politics and resources of research.

The human activity systems perspective can contextualise the current approach to ICT use in research.

Task Analysis and Design

Task analysis is a key technique of the Information Systems practitioner. This type of analysis identifies the task inputs, processes and outputs and how ICT can support the task.

Task analysis covers the traditional modeling of data and rules, but as importantly recognises a wide range of factors, including, at the *Process* level - actor knowledge & attitudes, procedures & policies that need to be complied with, tools & techniques used to complete the task, collaboration with colleagues, history & precedent, and situational context such as the organisation and power structure, degree of actor autonomy and so on. At the input/output, or *product*, level are the documents and data about the particular case to be processed and the deliverables produced by the process. These deliverables are the inputs to down-stream tasks.

The Process-Product model is complimented by 2 other levels, task management and task quality. *Quality* concerns (a) the attributes of the process that lead to effective deliverables and organisational learning, and (b) the attributes of the deliverables, in particular the use and impacts of the deliverables on all who will be impacted by them.

Task management includes task *planning* recognising the motivation, purpose and objectives of the task, risks, and the resources such as time and finances; task *monitoring* of financial and human resource use, progress reporting, etc and task *review*. These processes are, or should be, supported by traditional organisational information systems within the sponsoring university or research organisation.

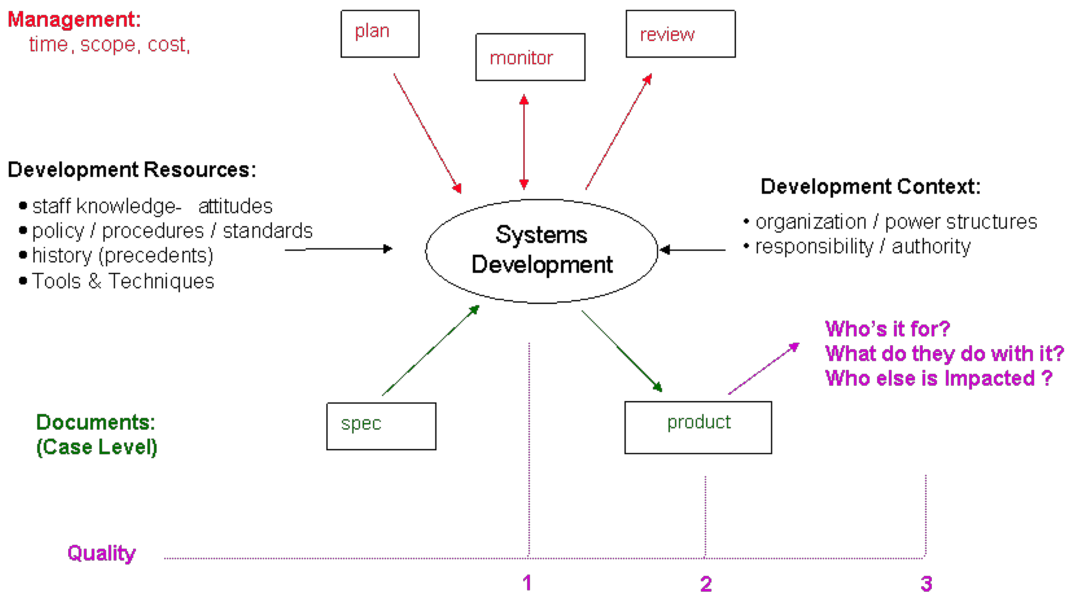

Figure 1. A general model of task analysis: the Product/Process

Of course, tasks can be conceived at micro level, such as a small data processing task, or at a process, sub-system, systems or enterprise level. tasks can be purely operational and repetitive, one-off tasks like the Sydney Opera House, or the majority of tasks that are in-between.

The wide variety of ICT facilities can be designed to support task execution. These include transaction processing systems, document and records management systems at the product level; knowledge management systems, information retrieval, simulations, expert systems and so on at the process level; project and process control systems at the management level and QA systems to support the quality aspects.

Domain Analysis

Different knowledge domains have very different characteristics and issues, but all use similar sets of IS approaches to systems conception, modelling, evaluation and so on.

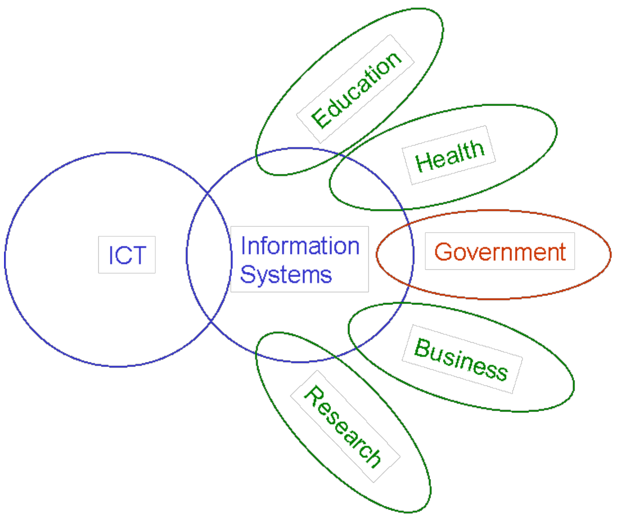

Figure 2. Information Systems linking ICT to various Knowledge Domains

ICT-based Information and Service Architecture

The perspectives noted above are not isolated. They need to be integrated into a whole to support a wide range of tasks in a coherent manner. This integrated architecture contrasts sharply with the infrastructure approach of the current e-research approach.

So what sorts of IS perspectives might usefully be considered for the Research domain?

E-RESEARCH AND INFORMATION SYSTEMS

A way to address this question is to consider current 'systems failures' in research tasks. Take a researcher who has some ill-formed questions and ideas about a problem area and starts to read around the topic. McDonald (2003) writing about the literature concerning disease in grape vines reported that it was dispersed, dated, under-utilised, expanding rapidly, variable in quality, variable in applicability, inconsistent, incomplete and slow to be published and applied. Clearly, there is a large problem to be addressed here even if the information management systems (document collection, indexing, bibliographic and full-text databases which store and deliver papers) were effective.

Consider an academic examining a PhD thesis. The literature problems above are similar, but in addition, the document is not usually available electronically, the data sets that were collected and analysed are not usually warehoused in an accessible way by the supervising university and the software used in the analysis is not always available. This issue is well recognised, for example by the UK's Digital Curation Centre: "data have importance as the evidential base for scholarly conclusions, and for the validation of those conclusions, a basic tenet of which is reproducibility. Curation is the active management and appraisal of data over the life-cycle of scholarly and scientific interest; it is the key to reproducibility and re-use".

There is a large population of people engaged in the Research domain - research funding bodies, research managers, the public, policy makers, etc. They all face difficult information problems.

In an Information Systems (IS) approach to these problems would need to deploy IS theories, tools and techniques will need to be deployed, reviewed and, probably, new IS approached developed. Some parts of an IS approach to e-Research might include:

research data warehouses and data mining
ontological systems for content organisation
meta-analysis to bring together work with similar ontological basis
more advanced techniques of domain analysis
knowledge management mechanisms to support research methods
serious e-libraries (for example DSpace)
agent technologies for search and evaluation
simulation and pattern matching facilities
development of domain-specific patterns
and so on.

But the issue is not the specific tools or technologies; it is how these might be systematically integrated into an ongoing socio-technical system.

Further, consider a professional who has the responsibility to make use of the latest research in her practice. We are stuck in a very outmoded system that serves neither researchers nor practitioners. There are attempts to address these problems. The Cochrane Collaboration has successfully adopted "systematic reviews" or meta-analyses as a method for getting best scientific results to practitioners and other researchers.

Knowledge Management Systems (KMS) technologies may be at the heart of a new kind of system that would be charged with representing the knowledge reported in a domain of research, and through a set of interface systems, employ the knowledge base in different ways to meet some of the needs in a range human activity systems. For example, a decision support system would use the KMS as a model of a domain to allow scenario processing; an expert system would give advice using the KMS as a knowledge base and justify the advice on the basis of the publications from which the KMS has been built; a Computer Aided Instruction (CAI) interface would allow the KMS to form the basis of courses in the domain; researchers and research bodies can use the KMS as a source for literature reviews and hypothesis testing. Each of these interface systems will have specific systems components suitable to their purposes but would rely on the core KMS as the source for their domain knowledge. The KMS would be self maintaining as each new research report that became available would be represented as a new document-related knowledge base and so participate immediately in the various uses to which the system is being put. Such a system would be domain specific, rather like the 'specialist libraries' of the past. The various needs of the different stakeholders could be met from a single core of knowledgebase, in the same way that previous generation systems used a single core database.

Proposals like this are not new. A century ago Paul Otlet was presenting a similar notion (see various papers by W.Boyd Rayward) and half a century ago it underlay Vannevar Bush's Memex system (1945). New technologies are being brought to bear on e-Research and IS should take a major role for without IS, another technology failure is imminent.

e-Research and e-Practice Inter-Operability

The KMS above introduces the second e-research issue that Information Systems can effectively address - the interaction of the research and practice human activity systems.

The idea that research results need to be socially useful is not new. The nature of knowledge production and use has long been a topic of debate and of academic research. In Australia, research has been largely a publicly funded activity, and Government is now casting an increasingly critical eye over the way it is currently performed. The Australian Research Council's *research network* initiative is the latest in a series of moves to promote inter-disciplinary research that aims to create and apply knowledge to address problems of national significance. Ronayne (1997) identifies two modes of research:

In Mode 1 problems are set and solved in a context governed by the interests of a largely academic community. By contrast, Mode 2 knowledge production is carried out within the context of application. It is intended to be useful to someone other than specifically the practitioner, be this industry, government or society generally; and this requirement is present from the beginning.

Batterham (2003) describes mode 1 as "discipline based; distinguishes between theoretical core and its conversion to application" while Mode 2 is "multidisciplinary, team based; Constant flow between basic & applied; Discovery occurs where knowledge is developed and put to use". The need for integration between research and practice is a government priority.

But research and practice are two very different worlds - human activity systems that share knowledge, but not purpose, method or people. The relationship has normally been one of information provision on one side and adoption on the other. There are many examples of effective inter-operation projects (for example, in most of Australia's Cooperative Research Centres. In this way of thinking the relationship between research and practice is like that between two organisations engaged in e-Commerce. Such inter-operation between systems has become a significant topic of IS research. Recent work in terms of ontologies and the web (Moody, 2003) and in object orientation, agents, XML, and so on is accelerating the field. So patterns to support this kind of relationship between research and practice already exist.

But the KMS described earlier implies a different view. It concentrates not on the trading between systems but on sharing what they have in common. It suggests a deeper association of human activity systems, not just ITC mediated communication. Figure 2 shows a very simplified view of a KMS as an information system mediating between the research and practice human activity systems.

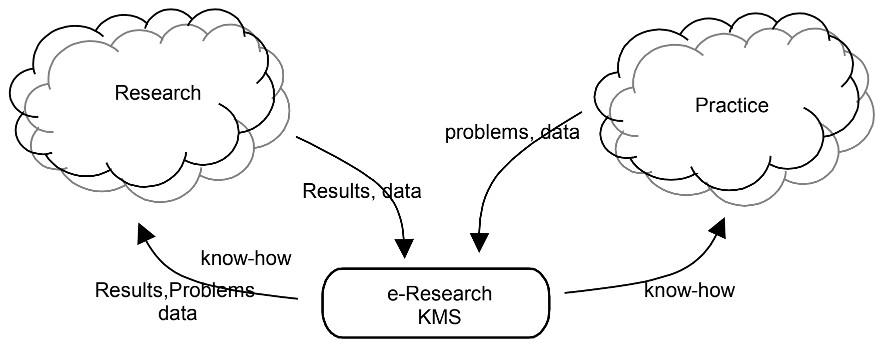

Figure 3: Inter-operation of Research and Practice systems (from McDonald 2004)

The effective component is knowledge (know-how) - with Practice providing relevance and raw data to used by Research, and Research providing economically useful causal knowledge and interventions to Practice. Systems, not technology, are key. Systems are instrumental in knowledge creation, knowledge management and knowledge mobilisation.

E-RESEARCH: CURRENT LIMITATIONS AND OPPORTUNITIES FOR IS

Most e-Research effort is currently focused on infrastructure building and on tools development. In the short term this will have pay-offs in large data-intensive research domains and in opportunistic niche uses of the technology. The coordination of ICT use by big science is very welcome.

But e-Research is much more that that. It is increasingly presented as a revolution in the way research is conducted. The sustainable value of the e-Research revolution it will only be realized by the broad-scale use of systems that employ that infrastructure. Such use will take many decades to develop if we only have a 'technology push' approach to e-Research.

A balance needs to be struck between the *human system* of research and the technology it can usefully deploy in the achievement of its ends. Models of technology maturation point to the demise of technology that is not built into the work practices of the domain it is meant to serve.

The co-ordination of e-Research infrastructure development needs to be complemented by a serious analysis of the human system of research and this analysis used to identify those parts of the system which would benefit most from ICT infrastructure support. These parts will not necessarily be obvious.

For example, e-Research is currently focused on data management (tools for data collection, storage, transmission and analysis) but what of the knowledge environment surrounding that data manipulation? There is little point in having cyber-age data processing with stone-age knowledge systems to manage ontologies, theories, previous work deployment of results etc.

The value of research is not in its existence but in the use of its results. e-Research needs to be complemented by e-Practice; the use of ICT to represent and deploy in industry and society the knowledge created by e-research.

An analysis of the human research domain would identify projects that would address systems-wide issues and change management processes for the broad research community. Such an approach contrasts sharply with technology-push approach which is beset with language like 'barriers to adoption' and 'how do we get buy-in?'.

The ICT noted in the levels above is used piecemeal by the Research human activity. At the foundation of the IS approach however are the very significant human, social and use aspects of technologies. Information Systems has the role of systematizing a range of ICT technologies that, combined, address a problem more effectively than individual technologies can. Technology is an *instrumental* component of the Information Systems approach to human activity systems, one that offers opportunities and limitations. At the moment, in the Research human activity system there is plenty of 'technology-push' in terms of new software and hardware, but little 'systems-pull' in terms of understanding and improving research operationally.

From an Information Systems perspective the Research domain presents an opportunity to develop and test new ideas in IS. Our challenge is to make a serious effort to create systems architectures, define and mobilize technologies and specify processes that address e-Research.

ICT does not work effectively in human activity systems without Information Systems. It does work piecemeal at the tools level for specific tasks but not for on-going management of data, information and knowledge; and not at the organisational or societal levels. So there is a demand and supply relationship between ICT and IS. From the IS perspective, new ICT offers opportunities for building new types of systems and it can be argued that IS has been slow to adopt some of these technologies. However, IS has certainly been slow in demanding the ICT tools needed to address new types of IS problem. Looking at Research as an IS domain challenges us to think of, and to start driving, the next steps in the development of IS into a comprehensive informatics discipline.

There are a couple of reasons why Research might not have so far been considered a domain of interest to IS. Firstly, research is usually seen as a project-oriented rather than an on-going activity. In a project, resources are assembled and used then disbanded when the project ends leaving little trace of its existence beyond its final report. Secondly, Information Systems usually concerns itself with organisational processes rather than with overall domains. This rather restricted focus has effectively excluded consideration of very significant issues which are now being accommodated both by alternate fields of study, notably Knowledge Management, and by alternate discipline-oriented approaches, for example Health Informatics (Health Informatics Society of Australia), bioinformatics, e-Learning, e-Government. These apparent alternatives ought to be IS application areas not competitors.

REFERENCES

Australian Partnership for Advanced Computing (APAC)
http://www.apac.edu.au/

AusGrid
http://www.ausgrid.org/

Batterham, Robin (2003) "E-Science: A Frontier Technology For Achieving the National Research Priorities" http://www.glassearth.com/seegrid_pdf/seegrid_batterham.pdf (accessed 30 September 2005)

Bauer, Henry (1992) *Scientific Literacy and the Myth of the Scientific Method* University of Illinois Press.

Bush, Vannevar (1945) "As we May Think" *Atlantic Monthly* June.

Checkland, P. & Scholes, J. (1990) *Soft Systems Methodology in Action* John Wiley.

Cochrane Collaboration
http://www.cochrane.org/

Cram, Laurence (2003) "A Roadmap for e-research" http://conferences.alia.org.au/alia2002/papers/cram_files/v3_document.htm (accessed 30 September 2005)

DEST (2005) Terms of Reference for the e-Research Coordinating Committee http://www.dest.gov.au/sectors/research_sector/policies_issues_reviews/key_issues/e_research_consult/tor.htm (accessed 30 September 2005)

Digital Curation Centre
http://www.dcc.ac.uk/

DSpace institutional repository system
http://www.dspace.org/

e-Social Science: Enabling Technologies and new Research Proceedings of Symposium, ANU, 12 November 2004.

Health Informatics Society of Australia
http//www.hisa.org.au

McDonald, Craig (2003) "eScience and ePractice: Reflections on a KMS development" *Proceedings of Australian Conference on Knowledge Management and Intelligent Decision Support*, Melbourne, Dec 2003.

McDonald, Craig (2004) "Research as an Information Systems Domain" *Proceedings of Information Systems Foundations Workshop*, ANU.

Moody, D (2000) "Building Links Between IS Research and Professional Practice: Improving the Relevance and Impact of IS Research" *International Conference on Information Systems* 2000.

National E-Science Center
http://www.nesc.ac.uk/

Netherlands Institute for Scientific Information Services
http://www.niwi.knaw.nl/en/the_virtual_knowledge_studio/

Rayward, W. Boyd (1991) "The Case of Paul Otlet, Pioneer of Information Science, Internationalist, Visionary: Reflections on Biography," *Journal of Librarianship and Information Science* 23 (September 1991)

Ronayne, Jarlath (1997) Research and the new universities: towards mode 2 *ASTE Focus* #98 http://www.atse.org.au/publications/focus/focus-ronayne.htm

Tsoi, A C (2004) "A Pilot in e-Research Support Discussion Paper" http://www.arc.gov.au/pdf/e-Research_support_discussion_paper.pdf

UK e-Science Grid
http://www.nesc.ac.uk/action/projects/project_action.cfm?Title=7 (accessed 30 September 2005)

ACKNOWLEDGEMENTS

This paper uses text from McDonald (2004) where necessary to fill in previously expressed principles and arguments where they are necessary to appreciate current developments in e-Research.